Teleform™ Scannable Data Entry: An Efficient Method to Update a Community-based Medical Record?

Paula Guerette, Ph.D., Bonnie Robinson, William P. Moran, M.D., Catherine Messick, M.D., Mark Wright, James Wofford, M.D., Ramon Velez, M.D., and the Community Care Coordination Network Database Group, Reynolds Health Center, Forsyth County NC, and the Department of Internal Medicine, Bowman Gray School of Medicine of Wake Forest University, and Wright Systems Solutions, Inc., Winston-Salem, North Carolina

Community-based multi-disciplinary care of chronically ill individuals frequently requires the efforts of several agencies and organizations. The Community Care Coordination Network (CCCN) is an effort to establish a community-based clinical database and electronic communication system to facilitate the exchange of pertinent patient data among primary care, community-based and hospital-based providers. In developing a primary care based electronic record, a method is needed to update records from the field or remote sites and agencies and yet maintain data quality. Scannable data entry with fixed fields, optical character recognition and verification was compared to traditional keyboard data entry to determine the relative efficiency of each method in updating the CCCN database.

INTRODUCTION

Secular trends, including a rapidly aging society and rising health care costs, have forced changes in the way health care is delivered; technological advances have further promoted these changes. For people with chronic health conditions, more agencies are involved in their care, and more people are receiving care at home. For example, it is not uncommon for an individual with a disabling condition such as a cerebrovascular accident to have a primary care physician, home health professionals such as nurse or therapist, and community-based services such as "meals-on-wheels" or respite care. Medical complications may force this individual to seek treatment in an emergency department or even be admitted to the hospital. To achieve a 'seamless' health care delivery system, the need for clinical information at these geographically (and usually organizationally) separate entities is crucial. Unfortunately, providers at different facilities must attempt to recompile the needed history, or provide care without the benefit of potentially useful patient information. Thus innovative information systems and communication links are needed to facilitate the timely, effective and efficient delivery of health care and services to patients living in the community[1].

One method to facilitate provision of care is to provide access to computerized data at numerous geographically separate locations. Since approximately 80% of older adults can identify their primary care provider [2], the primary care physician plays a significant role in the primary collection and maintenance of clinical data. The primary care provider is critical to decisions relating to the need for hospitalization and also authorizes most formal home health and community care services. Because of this, the primary care provider is ideally situated to develop and maintain a patient database.

The goals of the Community Care Coordination Network (CCCN) are as follows: 1) to identify functionally impaired chronically ill and disabled patients using a primary care clinical database, 2) to create a multi-disciplinary database for health care and community service personnel providing care to chronically ill and disabled patients, and 3) to facilitate the communication of changes in

0195-4210/95/$5.00

patient clinical and functional status among the primary care clinicians, health and community service agencies and the two county hospitals to assure critically important health and community care follow-up as these patients make transitions along the health care continuum. In-home data collection for the CCCN database requires approximately two hours per patient. However, in a population such as that included in the CCCN, patient status is constantly changing and some of the data can quickly become outdated. For example, it is not uncommon for a patient's functional status relating to daily activities to change. However, unless this patient visits the primary care physician, this change is frequently not reflected in the database. It is, nevertheless, critical to the validity of the network and to the quality of patient care that information be accurate and up-to-date. A major difficulty in maintaining such a database is the time and cost involved in data collection and entry.

Scannable data, in the form of mark-sense scanning to an exportable file format, is one potential means for populating and updating clinical databases. Scannable data on mark sense forms has been used to develop hospital-based [3] and community-based [4] databases. While scannable data entry requires an initial investment (i.e., cost of developing forms, cost of scanner), it may provide a cost-effective, timely, reliable method for creating and maintaining a clinical database. In addition, scanning techniques, through the use of fax-modem technology, may provide a feasible mechanism for updating a centralized clinical database from remote locations. The current pilot study was designed to compare the efficiency of scannable data entry with manual data entry.

METHODS

The Reynolds Health Center (RHC) provides primary care services to over 30,000 poor and medically indigent adults in Forsyth County, North Carolina. The CCCN at RHC provides a conduit for information sharing among several service providers in Forsyth County. The CCCN links two hospital emergency and social work departments, the Department of Social Services, Senior Services (a private, non-profit community service agency that assists functionally impaired elderly and administers a Medicaid waver program) and four home health agencies with the primary care providers at RHC. Patients with chronic conditions were identified from an existing RHC patient database and were selected for inclusion in the CCCN project. These patients were recruited because they are typically high users of a variety of community and hospital-based health care services. An array of information, including patient cognitive, functional and affective status, home environment, services received, and caregiver support [2] was collected in the field for inclusion in the CCCN database. Previously standardized and validated data collection instruments were used when available. Detailed information in the database is currently available for over 600 patients with chronic health conditions.

Representatives from each of the agencies in the CCCN met over a period of nine months to establish a consensus of the types of data that should be included in the CCCN database. This was done for two purposes: 1) to promote the utility of the database for network members, and 2) to provide a consensus on form content to allow remote data exchange at a later point in the project. The total CCCN assessment battery includes twelve domains, eleven of which were considered appropriate for scanning. The remaining form, a detailed functional assessment, included a majority of specialized information (e.g., caregiver names, addresses, etc.) which required constrained print field data entry for scanning. Since the optical character recognition (OCR) features of the constrained print fields used in scanning require careful, neatly printed data, it was determined that this format would severely slow the data collectors. The eleven forms that were created in scannable format required some constrained field printing, but included a majority of fixed field choices (e.g., gender,

race, yes/no questions).

A commercially available program from Cardiff Software [5] called *Teleform* ™ was used to develop the data collection forms. This program allows the simultaneous creation of a database structure when the forms are developed, eliminating the need to create a data entry form separately. Each form is automatically given a unique barcode form ID number and prints four 'cornerstones' to provide form orientation. Data scanned into this program are routed through a 'verifier' to compare the computer interpretations of the scanned responses with various constraints specified by the form developer (e.g., 'single response choice required,' 'response required always'); constraints can be tailored for each item and serve to improve accuracy of scanning. Once the data have been scanned and verified, they can be automatically or manually exported into the computerized medical record in various formats, including delimited text, comma separated values and a variety of common database (e.g., dBase, FoxPro, Paradox) or spreadsheet (e.g., Lotus 123) formats.

Two representative forms out of the twelve in the CCCN assessment battery were selected for use in the pilot study. One form was comprised primarily of fixed choice fields. This form was representative of the majority of forms in the overall battery. The other form included a combination of constrained print fields, and forced choice responses. To compare scanning and manual data entry efficiency, forty sets of forms were randomly chosen out of the 600 CCCN data sets. The two selected forms from twenty sets were scanned while these same forms from the other twenty were entered manually. In addition, data entry for the *entire* CCCN data set was completed (manually entered $n=441$, scanned $n=166$) and allowed a subsequent review of the error rates for the constrained print field form across the entire data set.

RESULTS

Accuracy. Independent groups t-tests were performed on the accuracy of data entry for the two methods (scanning and manual entry) for the two different form types. There was no significant difference in the accuracy of data across the two data entry methods for either form type. Also, in the subsequent review of the error rates for the constrained print form across the entire data set, it was found that the error rates for the scanned data were nearly three times that for the manually entered data. However, error rates for both data entry methods were extremely low (manual - 0.5%, scanned - 1.4%), making this a trivial issue.

Speed. Independent groups t-tests were also performed on the speed of data entry for the two scanning methods for the two different form types. For the fixed choice data form, the total time for the scanning process was significantly slower than manual data entry (108.8 seconds per record as compared to 28.4 seconds; $t=31.9$, $p<.001$, $df=38$). However, the scanning data entry process is actually comprised of two stages: scanning and verification. The scanning stage is completely automated, while the verification stage requires human intervention. In comparing only the actual human time required for each process (i.e., verification only versus manual data entry), verification time is significantly shorter (17.6 seconds as compared to 28.4 seconds; $t=-4.1$, $p<.001$, $df=38$). Overall, the scanning stage comprised 84% of the total data entry time and for the constrained print form, and 52% of the time for the fixed choice form.

For the constrained print field data form, there was no significant difference in total time required to scan versus to enter data manually. When the verification time was compared to the manual data entry time, there was a significant difference. The time to verify the scanned from was significantly less (97.7 seconds as compared to 225.0 seconds; $t= -9.3$, $p<.001$, $df=38$) than the time needed for manual data entry.

DISCUSSION

When considering only the actual personnel time required to attend to and enter data, scannable data entry appears to be a time efficient means for inputting both fixed choice and constrained print data into a clinical database. The total time for scanning was, in the case of the fixed response choice form, longer than that of manual data entry. However, the actual personnel time required during the scanning process was only a small portion of the total time required (i.e., 16%), thus freeing the data processor up to do other tasks during the scanning stage of the electronic data entry.

Beyond this, there are interesting implications for updating data collected by community-based health providers. The scanning method used to populate part of the original CCCN database can also be used to update the data in an ongoing manner. Members of the network have determined the types of data currently most useful to collect for the CCCN database, and those data of potential use in the future to support clinical decisions. Some of these data (e.g., functional assessments, cognitive assessments) are collected in an ongoing fashion in the field by home health professionals. By using the scannable forms, these professionals can update the CCCN database simply by faxing the forms to the CCCN. Data received from the CCCN fax can be automatically exported into the database after verification, providing a chronology of information about each domain and allowing other providers to benefit from the most current information. This is also a cost effective approach to sharing data which will often preclude the need to recollect similar data at various locations. This automated approach also provides a mechanism for insuring the quality of the data while still allowing bidirectional data updating. The feasibility for updating the CCCN database by fax with data gathered by community-based providers is currently being explored.

Several caveats with the scanning process should be addressed however. One particular difficulty was apparent with the OCR process during verification of the constrained print fields. These types of data fields require extreme precision in data form completion. Letters must be neatly printed in black ink and may not touch the edge of the constrained field boxes in order to be read properly by the scanner. This slowed down the data collection process, and caused difficulty with OCR with some forms that were not neatly printed. In addition, fixed choice fields that were completed in pencil or by using a check mark or an "x" were frequently misread by the scanner. Thus, careful, extensive training of interviewers is required in order to insure the readability of the forms. This could cause some difficulties with forms completed at remote sites by clinicians with less familiarity with the form completion requirements.

Mechanical difficulties were encountered with the scanner, such as pulling through multiple pages simultaneously. If a page was pulled into the scanner even slightly too far, the verifier would not recognize the cornerstones and would indicate that the form was incomplete, necessitating rescanning and potentially adding significantly to the time required. In the pilot study, batches of five forms were used in making estimates of the time required for scanning, and larger batches would reduce the per form verification time. However, in reviewing this issue with the overall CCCN forms (in which larger batches were used), it was discovered that entire forms *frequently* required rescanning due to difficulty reading one or more data fields. This resulted in significant frustration and time lost. The use of a "page link" feature provided in *Teleform* may allow rescanning of only pages that were misread, rather than entire forms.

Finally, extensive difficulties were encountered in attempting to export the data. For example, several field names were inadvertently assigned which were "SQL Keywords" (e.g., date, year, money). This caused multiple problems, including fatal application errors

and loss of data. Lack of adequate technical support further contributes to ongoing difficulties with scannable data entry and export.

In conclusion, optical scanning using a combination of optical character recognition and mark-sense reading offers the *potential* for efficient clinical data entry to a primary care clinical database. However, significant care must be taken in data form development, completion, scanning and exporting using *Teleform*. The use of standardized clinical data instruments and scannable data entry may offer another method of data entry for community-based care of chronically ill individuals.

References

[1] Pope AM, Tarlov AR (eds). Disability in America: Toward a National Agenda For Prevention. Washington, DC: National Academy Press, 1991.

[2] Moran WP, Messick C, Guerette P, Anderson R, Bradham D, Wofford JL, Velez R. A practice-based information system for multi-disciplinary care of chronically ill patients: What information do we need? 1994; Proceedings of the 18th Annual Symposium on Computer Applications in Medical Care, Washington, DC, 1994:585-589.

[3] Lindemann M, Rock P, Locke JA, Johnson R. Optical mark reader technology and data base development in anesthesiology and critical care medicine and the Johns Hopkins Medical Institutions. In Proceedings of the 15th Annual Symposium on Computer Applications in Medical Care, Washington, DC, 1991:798-802.

[4] Moran WP, Wofford JL, Hamrick V, Myers BT, Cooper J, Doby J, FitzGerald K, Spence C & Velez R. Implementing computerized tracking at a community health center: Challenges and solutions. In Proceedings of the 17th Annual Symposium on Computer Applications in Medical Care, Washington, DC, 1993:139-143.

[5] Teleform 4.0 for Windows, Cardiff Software, Inc., Solana Beach, CA, 1991-1994.